Supporting Information

Engineering a Rigid Protein Tunnel for Biomolecular Detection

Mohammad M. Mohammad[1], Raghuvaran Iyer[2], Khalil R. Howard[3], Mark P. McPike[4], Philip N. Borer[2,3,4] and Liviu Movileanu[1,3,5]

[1]*Department of Physics, Syracuse University, 201 Physics Building, Syracuse, New York 13244-1130, USA*

[2]*Department of Chemistry, Syracuse University, Syracuse, New York 13244-4100, USA*

[3]*Structural Biology, Biochemistry, and Biophysics Program, Syracuse University, 111 College Place, Syracuse, New York 13244-4100, USA*

[4]*AptaMatrix, Inc., 100 Intrepid Lane Suite 1, Syracuse, NY 13205, USA*

[5]*The Syracuse Biomaterials Institute, Syracuse University, 121 Link Hall, Syracuse, New York 13244, USA*

Running title: Protein engineering for biomolecular detection

Correspondence/reprint requests:
Liviu Movileanu, PhD, Department of Physics, Syracuse University, 201 Physics Building, Syracuse, New York 13244-1130, USA; Phone:315-443-8078; Fax: 315-443-9103;
E-mail: lmovilea@physics.syr.edu

Preparation of the *fhua Δc/Δ4l* gene. The construction of the *fhua Δc/Δ4l* gene and its cloning into the expression vector have been reported previously [1]. Briefly, the *fhua* gene lacking the loops *3, 4, 5, 11* and the first 160 amino acids, named *fhua* Δ *cork Δ 4 loops* (*fhua Δc/Δ4l),* was constructed by *de novo* synthesis (Geneart, Regensburg, Germany) in the pMK-RQ plasmid flanked by EcoR*I* and Xho*I* restriction sites for cloning purposes. In this construct, the deleted loops were replaced with NSEGS polypeptide linker[2]. The pMK-RQ plasmid was digested with the EcoR*I* and Xho*I* enzymes, and the released *fhua Δc/Δ4l* gene was gel purified using the Minielute gel purification kit (Qiagen, Germantown, MD) and cloned into the pPR-IBA1 expression vector pre-digested with EcoR*I* and Xho*I* enzymes. A C-terminal 6X Histidine (6×His$^+$) tag was added to *fhua Δc/Δ4l* by overlap extension PCR to produce the final expression vector (pPR-IBA1- *fhua Δc/Δ4l*-6×*His*$^+$). The final plasmid was verified by DNA sequencing.

Overexpression and purification of the FhuA ΔC/Δ4L protein under denaturing condition. The protein was expressed in *E. coli* BL21 (DE3). Cells, transformed with pPR-IBA1- *fhua ΔC/Δ4l-His*$^+$ plasmid, were grown in *2X* TY media at 37°C until OD_{578} ~0.7-0.8, at which time the expression of the FhuA ΔC/Δ4L protein was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Expression was allowed to continue until the cell growth plateaued, as measured by OD_{578}. Then, cells were harvested by centrifugation and the pellet was resuspended in resuspension buffer (100 mM NaCl, 50 mM Tris-Cl, 10 mM $MgCl_2$, pH 8.0) supplemented with 10 µg/mL DNase I and protease inhibitors (Roche, Branchburg, NJ). The resuspended cells were lysed using a Microfluidizer, model 110L (Microfluidics, Newton, MA). Following cell lysis, the homogenate was centrifuged for 20 min at 16 000 *xg*, 4°C. Pellets containing inclusion bodies were resuspended in inclusion body-cleaning buffer (50 mM Tris-HCl, 0.1% Triton X100, pH 8.0), homogenized using Potter-Elvehjem homogenizer (VWR, Bridgeport, NJ), and recentrifuged for 20 min at 16 000 *xg*. This step was repeated three times. The final pellet was resuspended in denaturing buffer (50 mM Tris-HCl, 8 M urea, pH 8.0) at the concentration of ~5 mg/ml and then recentrifuged to remove the insoluble materials. The final FhuA ΔC/Δ4L-containing solution was filtered through 0.2 µM filters (Thermo Fisher Scientific, Rochester, NY). The solubilized FhuA ΔC/Δ4L was loaded onto a column packed with 2 ml of Ni$^+$-NTA resin (Bio-Rad, Hercules, CA), which was equilibrated in 500 mM NaCl, 50 mM Tris-HCl, 8 M urea, pH 8.0. The column was washed in two steps with 5 and 25 mM imidazole in the same equilibrating buffer. The FhuA ΔC/Δ4L protein was eluted with equilibrating buffer containing 350 mM imidazole in 5 ml fractions. SDS-PAGE was used to monitor the elution profile of pure proteins.

Refolding of FhuA ΔC/Δ4L: Refolding of FhuA ΔC/Δ4L protein was executed by the protocol developed by Kleinschmidt and colleagues [3]. The refolded proteins were obtained by diluting denatured protein into solution that contained the desired detergent. 40 µl of purified, denatured FhuA ΔC/Δ4L was diluted 50-fold in dilution buffer (50 mM Tris-HCl, 1 mM EDTA, 200 mM NaCl, pH 8.0) containing 1.5% n-Dodecyl-β-D-maltopyranoside(DDM), 2.5% n-Octyl-β-D-Glucopyranoside(OG) or 16 mM 1-lauroyl-2-hydroxy-sn-glycero-3-phophocholine (LPhC). The mixture was incubated overnight at 23 °C to complete the refolding of FhuA ΔC/Δ4L proteins. The unfolded proteins or protein aggregates were removed by centrifugation.

Electrical recordings on planar lipid bilayers. Electrical recordings were carried out with planar bilayer lipid membranes (BLMs) [4]. The *cis* and *trans* chambers (1.5 mL each) of the apparatus were separated by a 25 µm-thick Teflon septum (Goodfellow Corporation, Malvern, PA). An aperture in the septum of 50 µm in diameter was pretreated with hexadecane (Sigma-Aldrich, St. Louis, MO) dissolved in highly purified pentane (Fisher HPLC grade, Fair Lawn, NJ) at a concentration of 10% (vol/vol). A 1,2 diphytanoyl-*sn*-glycero-phosphatidylcholine (Avanti Polar Lipids, Alabaster, AL) bilayer was formed across the aperture. The FhuA ΔC/Δ4L protein nanopores were added to a final protein

concentration of 100-180 ng/ml. Single-channel currents were recorded by using a patch clamp amplifier (Axopatch 200B, Axon Instruments, Foster City, CA) connected to Ag/AgCl electrodes through agarose bridges. The *cis* chamber was grounded so that a positive current (upward deflection) represents positive charge moving from the *trans* to *cis* side. A Precision T3500 Tower Workstation Desktop PC (Dell Computers, Austin, TX) was equipped with a DigiData 1322A A/D converter (Axon) for data acquisition. The signal was low-pass filtered with an 8-pole Bessel filter (Model 900; Frequency Devices, Ottawa, IL) at a frequency of 10 kHz and sampled at 50 kHz, unless otherwise stated. For data acquisition and analysis, we used the pClamp9.2 software package (Axon).

Barnase proteins: Barnase constructs (pb_2(35)-Ba, pb_2(65)-Ba and pb_2(95)-Ba) were prepared as previously described [4,5]. Briefly, pb_2-Ba containing plasmids were used to express pb_2-Ba proteins in *E-coli*. Proteins were denatured from inclusion bodies by 6 M guanidine hydrochloride, and were dialyzed against the dialysis buffer to refold the proteins (50 mM NaOAc/HOAc, 1.5 M guanidine hydrochloride, 0.5 mM PMSF, 2 mM EDTA, pH 5) for 3 h at 4 °C. Refolded proteins were diluted into 50 ml in dilution buffer (50 mM NaOAc/HOAc, 2 mM EDTA, 5 mM DTT, pH 5), and then concentrated (2-4 mg/l of cell culture).

Pepsin and mouse serum immunoglobulin G (IgG) proteins: Porcine pepsin (3802 units/mg) and the IgG (purity > 90%) were purchased from Sigma (St. Louis, MO). Purities of both proteins were in-house checked by SDS-PAGE (**Fig. S11**, lane 1 and lane 2).

NCp7 protein. NCp7 protein was prepared, as described previously [6,7]. Protein stocks at 50-100 μM were kept at -80°C for several months in 200 mM NaCl, 50 mM Tris, 1μM $ZnCl_2$, 10% glycerol, 1 mM tris(2 carboxyethyl) phosphine (TCEP), pH 8.0. Fluorescence titrations assays with SL3 added to 0.3 μM NCp7 were used as a quality control for NCp7 protein [6,7]. We routinely reject the NCp7 preparations that did not exhibit at least 95% quenching with a 3-fold excess of the 20mer, SL3 5`-GGACUAGCGGAGGCUAGUCC-3` (K_d is 21 nM) [8,9]. In case of single-channel electrical recordings, we used DNA 1 aptamer as quality control for NCp7 preparations and its aliquots. Aliquots of protein that were used for DNA 2 and DNA 3 titration experiments were checked with DNA 1 aptamer to assure that NCp7 was active. DNA 1 aptamer typically reduces the frequency of the currents events caused by the interaction of NCp7 with FhuA ΔC/Δ4L by 85-90 %, and aliquots that did not give this percentage were not used in the nanopore experiments.

Identification of high-affinity DNA aptamer for the NCp7 protein. The high-affinity aptamer for NCp7, which is shown in **Fig. 5a** (DNA 1), was identified within a library of all 4^6 possible hexanucleotide loops (M.P. McPike, J.E. Crill, P.N. Borer, *unpublished*). Briefly, each fixed-sequence stem and loop was synthesized on a microarray flowcell (Geniom One, Febit, Inc., Heidelberg, Germany) with one aptamer candidate per microarray feature. Lysine residues of the NCp7 protein were labeled with an average of one Cy3 fluorophore per protein, and flowed over the microarray. Relative affinities of aptamer candidates for NCp7 were ranked by fluorescence intensity. DNA 1 had high intensity and both DNA 2 and DNA 3 had low intensity.

DNA stem-loops. PAGE-purified DNA 1, 2 and 3 were purchased from IDT DNA Technologies (Coralville, IA). DNA samples were heated to 95°C for 5 minutes, cooled on ice and remained on ice until used in either fluorescence or nanopore-based titrations experiments. According to OligoAnalyzer 3.1 software from IDT, all three DNAs should fold as the base-paired stems and hairpin loops shown in main text **Fig. a** (the free energy of folding ($\Delta G°$) the DNA stem-loops from unpaired random coils -10.3, -11.5 and -9.8 kcal/mole for DNA 1, DNA 2 and DNA 3, respectively, in 0.2 M NaCl).

NCp7 titration assays. Fluorescence assays were performed by adding DNA stem-loops (50 µM stock) in aliquots of 1.5-3 µl to 0.30 µM NCp7 in 2.0 ml of the fluorescence buffer (5 mM sodium phosphate, 0.1% PEG, 0.2 M NaCl, 1 µM $ZnCl_2$, 1 mM TCEP, pH 7 in a 10 mm X 10 mm path length quartz fluorescence cuvette (NSG Precision Cells, Farmingdale, NY) along with an 8 mm X 1.5 mm stir bar. The fluorescence of the free NCp7 (the fraction of the protein that is not bound to the DNA) was measured using a PTI QM-4/2003SE fluorometer at 349.9 nm emission using a 290 nm excitation, a 4 nm excitation band-pass, and an 1.5-3 nm emission band-pass. In protein nanopore-based titrations assays, 1 µM NCp7 was added to a single FhuA ΔC/Δ4L nanopore inserted into the lipid bilayer in 1.5 ml chamber in the same titration buffer mentioned above. To obtain fraction of the protein that is not bound to the DNA, the frequency of the currents events caused by the addition of the NCp7 protein was calculated before and after the addition of DNA.

In both assays, the titration curves obtained were fitted to a model assuming 1:1 stoichiometry for the ratio of NCp7 bound to DNA (see the main text). The titration curves were fitted as previously [6]:

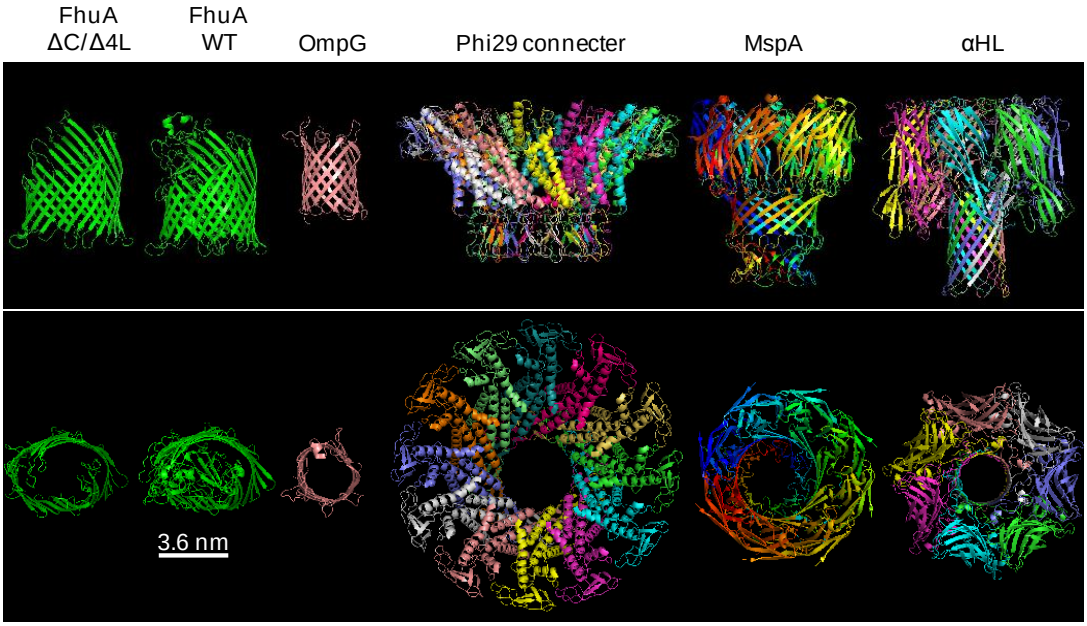
$$\frac{I - I_\infty}{I_0} = \frac{-\left(D_t - P_t + K_d\right) + \left[\left(D_t - P_t + K_d\right)^2 + 4P_tK_d\right]^{1/2}}{2P_t} \qquad (1)$$

where P_t and D_t are the total protein and DNA concentrations, respectively. I is the measured fluorescence intensity or the normalized frequency of current events. I_0 is the fluorescence intensity or the normalized event frequency when D_t=0. I_∞ is the limiting intensity at saturation or the normalized event frequency at saturation. K_d denotes the dissociation constant. The equation was derived using the following standard assumptions: (i) I - I_∞ is directly proportional to the free protein concentration; (ii) P_t=P_f + P_{bound}, where P_f is the free concentration of the NCp7 protein and P_{bound} is the concentration of the NCp7 protein bound to the DNA. D_t=D_f + D_{bound}, where D_f is the free concentration of the DNA and D_{bound} is the concentration of the DNA bound to the NCp7 protein; (iii) K_d = (D_fP_f)/(PD), where PD is the concentration of the NCp7-DNA complex. The data were fitted using Origin 8.1 software. P_t was fixed at 0.30 µM in the fitting of fluorescence titration data, and at 1 µM in the fitting of the nanopore titration data. I_∞ was fixed at zero in either case. K_d was the only adjustable parameter. There was no improvement in the fit when I_∞ was allowed to be an adjustable parameter along with K_d.

Calculating the ionic selectivity of FhuA ΔC/Δ4L. The ionic selectivity of the FhuA ΔC/Δ4L was determined by reconstituting FhuA ΔC/Δ4L single channels into a planar lipid bilayer in symmetrical solutions of 0.02 M KCl, 10 mM potassium phosphate, pH 7.4. Then, the *trans* chamber was offset to 0.2 M KCl, 10 mM potassium phosphate, pH 7.4. The single-channel current was then recorded versus the transmembrane potential to determine the reverse potential (V_r, **Fig. S5**). The reverse potential V_r is defined as the transmembrane potential corresponding to a zero single-channel current recorded in asymmetrical conditions. The permeability ratio P_K / P_{Cl} was calculated from the reverse potential (V_r) by applying the Goldman-Hodgkin-Katz equation:

$$\frac{P_K}{P_{Cl}} = \frac{\left[a_{Cl}\right]_t - \left[a_{Cl}\right]_c \exp\left(\frac{V_rF}{RT}\right)}{\left[a_K\right]_t \exp\left(\frac{V_rF}{RT}\right) - \left[a_K\right]_c} \qquad (2)$$

Where a, F, R and T are the activity coefficient, the Faraday constant, the gas constant and the absolute temperature, respectively.



	FhuA ΔC/Δ4L	FhuA WT	OmpG	Phi29 connecter	MspA	αHL
Pore diameter (nm)	3.1 x 4	3.1 x 4	1.4	3.6	1	1.4
Conductance (nS)	3.9	0.3	1.2	4.8	4.9	0.9
Number of subunits	1	1	1	12	8	7

Fig. S1. Crystal structures of protein nanopores used for single-molecule detection. The cartoon indicates the side (upper panel) and top (bottom panel) views. Multimeric protein nanopores are shown with subunits of different colors. The PDB coordinates, from right to left, are 7AHL [10], 1FOQ [11], 1UUN [12], 2F1C [13] and 1BY5 [14] for staphylococcal α-hemolysin (αHL), mycobacterium outer membrane protein A (MspA), membrane-adapted phi29 motor protein connecter, outer membrane protein G of *E. coli* (OmpG) and ferric hydroxamate uptake component A of *E. coli* (FhuA), respectively. The structural model of FhuA ΔC/Δ4L was based upon the crystal structure of the native FhuA protein [14].

Table S1: Comparison of the unitary conductance measured with FhuA ΔC/Δ4L refolded in different detergents.

Detergent	Number of independent experiments	Conductance (nS) [a,b]
n-dodecyl-β-D-maltopyranoside (DDM)	7	3.9 ± 0.5 (n=92)
n-octyl-β-D-glucopyranoside (OG)	8	3.8 ± 0.8 (n=33)
1-lauroyl-2-hydroxy-sn-glycero-3-phosphocholine (LPhC)	6	3.9 ± 0.4 (n=44)

[a]The conductance was obtained from the ratio of the measured current discrete steps to the applied transmembrane potential of +40 mV. The buffer solution was 1 M KCl, 10 mM potassium phosphate, pH 7.4.

[b]n represents the total number of discrete single-channel insertions.

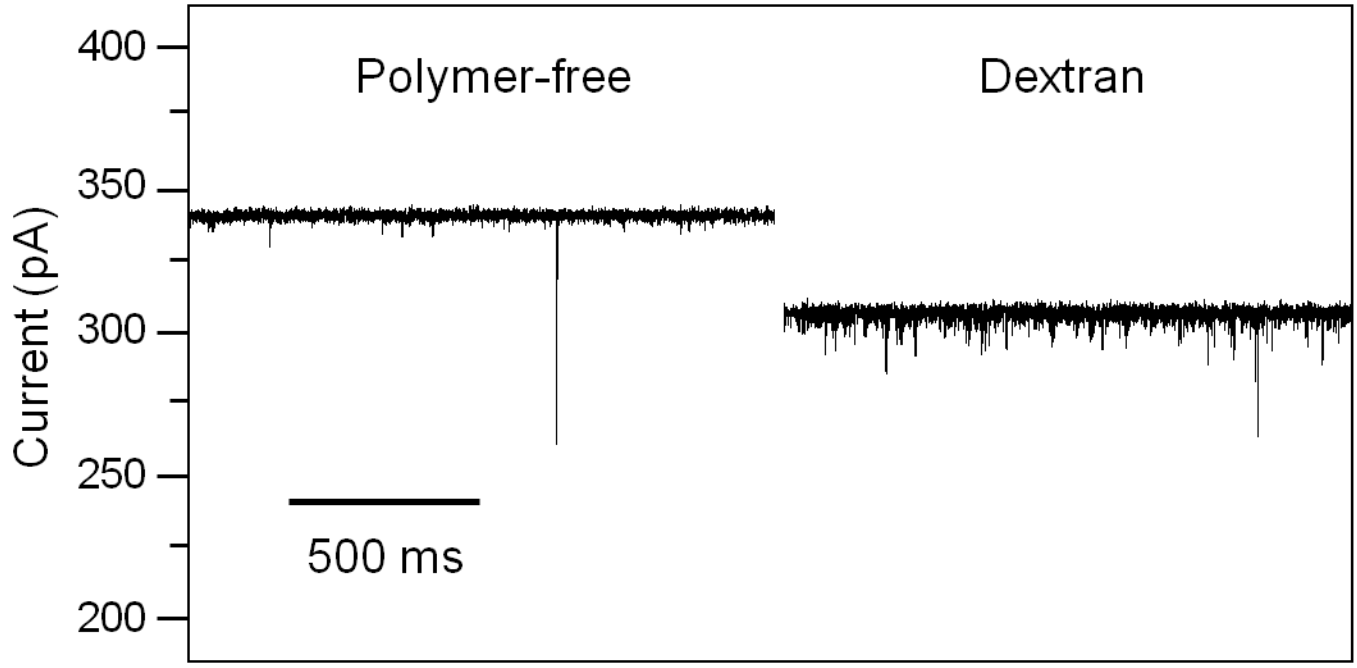


Fig. S2. 40-kDa dextran polymer alters the ion current of FhuA ΔC/Δ4L nanopore. Representative single-channel electrical recording shows the effect of dextran on the single-channel current of the FhuA ΔC/Δ4L protein nanopore. A reduction in the single-channel current of 9 % was observed after the addition of dextran to both sides of the chamber. Polymer concentration in the bulk was 15% (w/w) in 1M KCl, 10 mM potassium phosphate, pH 7.4, with an applied potential of +80 mV. The channel used in this experiment had a unitary conductance of 4.2 nS. Electrical traces were low-pass Bessel filtered at 2 kHz.

Dextran is a water-soluble, neutral polymer that has been used to estimate the access-resistance and radii of several protein ion channels [15,16]. Typically, the pore is modeled as a cylinder. The total estimated resistance of the pore is contributed by the intrinsic resistance of the pore itself and its access-resistance. In the case of a simple cylinder, the access-resistance can be expressed as: $R_a = 1/(2\,\sigma\, r)$, where σ is the conductivity of the solution and r is the cylinder radius [17].

By adding dextran to the bulk solution surrounding the nanopore, the conductivity, σ, is decreased (**Fig. S2**), changing the access-resistance. If dextran is large enough such that it does not enter into the nanopore interior, the intrinsic resistance of the pore is left unaltered. Following Bezrukov and Vodyanoy (1993), the radius of the channel may then be inferred using the equation

$$r = \frac{q^*}{2\sigma^*}\frac{1-\sigma^*/\sigma}{1-q^*/q}$$

where q and q^* are the measured conductance of the pore with and without dextran, respectively. σ and σ^* are the bulk conductivity of the solutions with and without dextran, respectively.

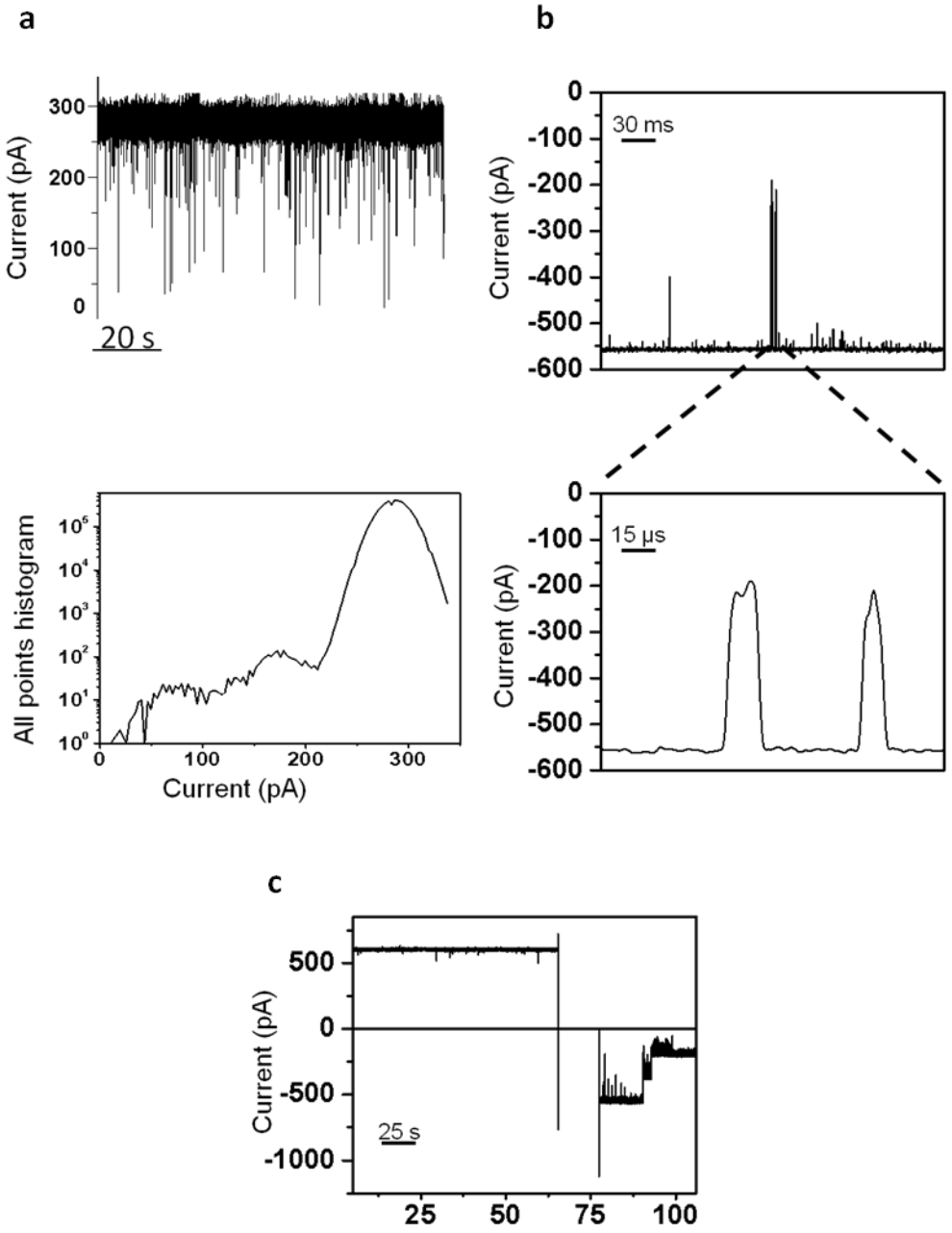


Fig. S3. FhuA ΔC/Δ4L nanopore inserts in the lipid bilayer as monomer into a single orientation. The transient single-channel current blockades made by long polypeptide chains were up to 95% of the unitary current of the engineered protein, suggesting that the nanopore inserts into a planar lipid bilayer as a monomer. This protein nanopore inserted into the synthetic bilayer in a single orientation. For example, at an applied transmembrane potential of 140 mV, the voltage-induced gating of the single-channel current was always dominant at a negative bias. **a**, Single-channel recording showing transient single-channel current blockades made by barnase protein fused to positively-charged presequence $pb_2(35)$ ($pb_2(35)$-Ba). 200 nM of $pb_2(35)$-Ba was added to *trans* side of the chamber. The lower panel presents all-points amplitude histogram of the above trace. **b**, Higher transmembrane potential produce closures of more than 50% of unitary current in the single-channel electrical recording. **c**,Typical single-channels recording showing that the FhuA ΔC/Δ4L nanopore has higher probability of channel closure at negative voltage. All experiments were carried out in 1 M KCl,10 mM

phosphate buffer with, pH 7.4, and at the applied potential of +80 mV, -140 mV and ±140 mV for **a**, **b** and **c**, respectively. Electrical traces low-pass Bessel filtered at 2 kHz.

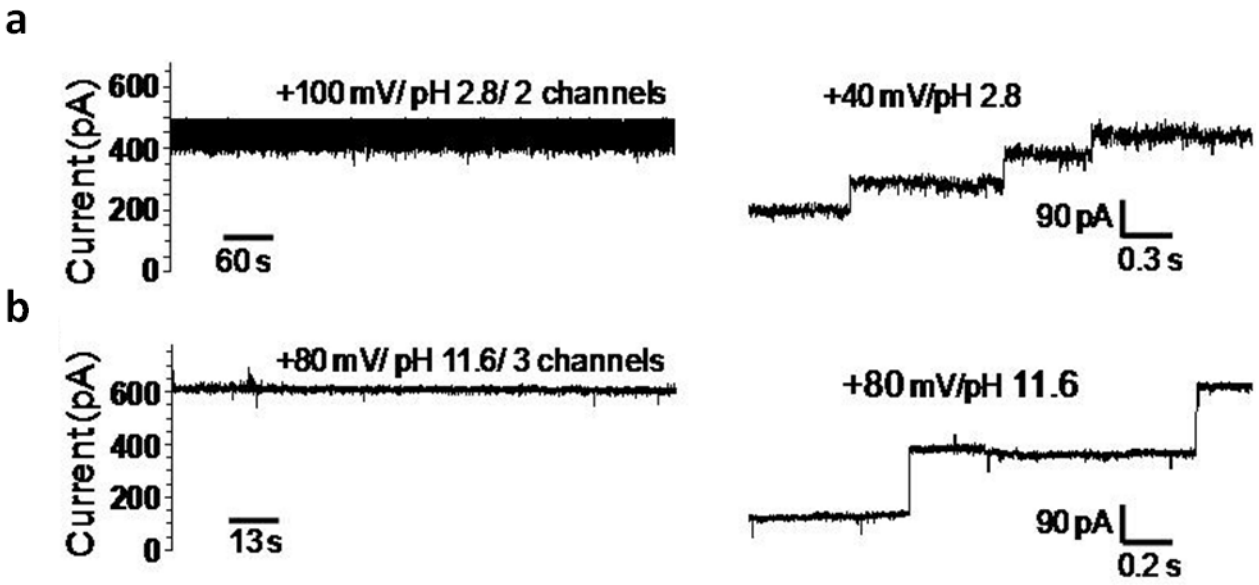


Fig. S4. FhuA ΔC/Δ4L protein nanopore is stable at extreme pH conditions. **a**, FhuA ΔC/Δ4L protein nanopore is stable at very low acidic pH. *Left panel* shows electrical recording with two channels at an applied potential of +100 mV. *Right panel* shows how FhuA ΔC/Δ4L protein nanopore inserts into a synthetic bilayer at highly acidic pH. The trace presents stepwise single-channel insertions at an applied potential of +40 mV. These experiments were performed in 1 M NaCl, citrate-buffer, pH 2.8. **b**, FhuA ΔC/Δ4L protein nanopore is stable at high alkaline pH. *Left panel* shows electrical recording with three channels at an applied potential of +80 mV. *Right panel* shows how FhuA ΔC/Δ4L protein nanopore inserts into synthetic bilayer at an alkaline pH. The trace indicates stepwise single-channel insertions at an applied potential of +80 mV. These experiments were carried out in 1 M NaCl, $Na_2HPO_47H_2O$ buffer, pH 11.6. Traces in **a** and **b** were low-pass Bessel filtered at 2 kHz.

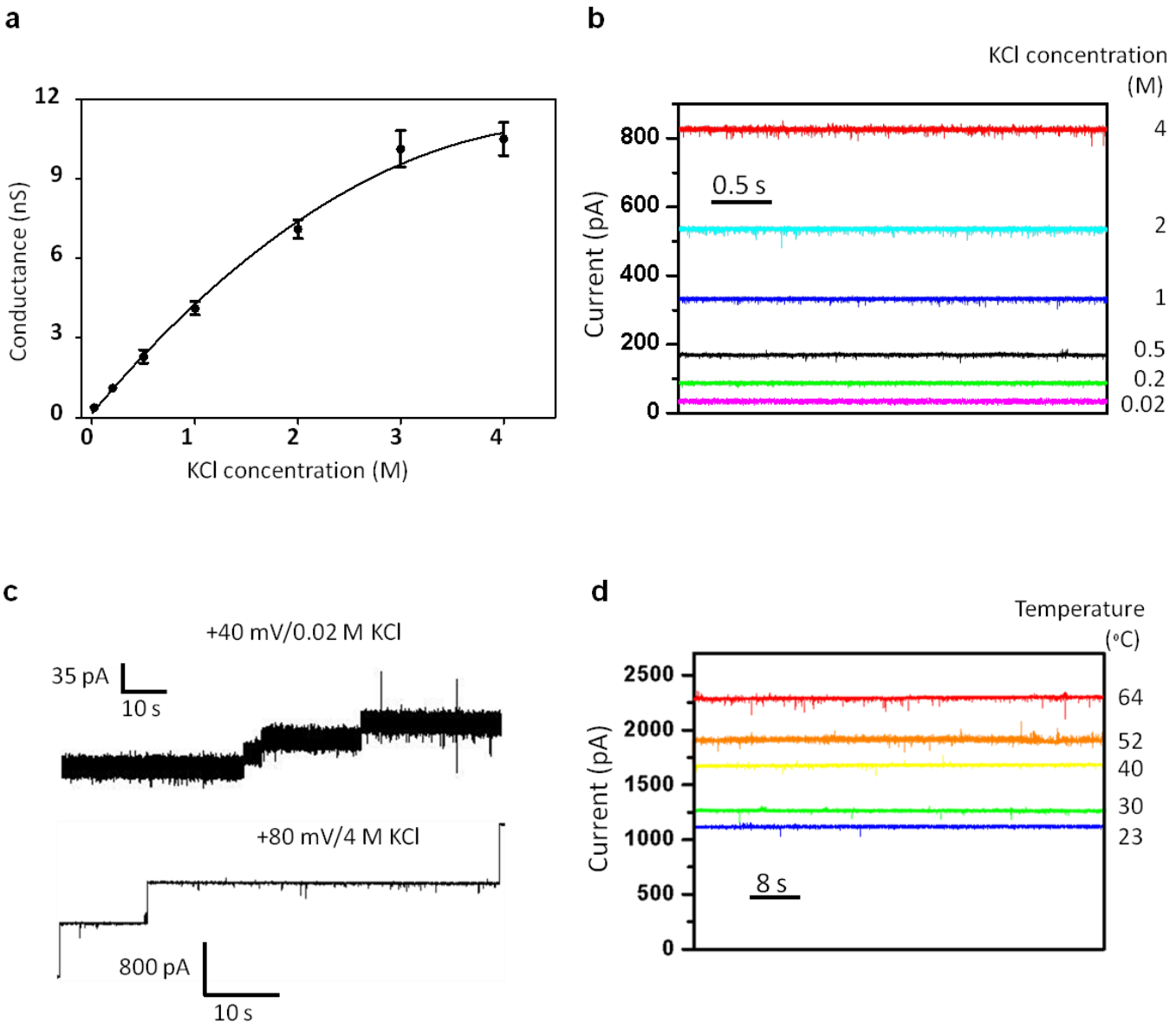


Fig. S5. FhuA ΔC/Δ4L protein forms a rigid and unusually rigid protein nanopore. a, Single-channel conductance of the FhuA ΔC/Δ4L protein nanopore as a function of the KCl concentration. At each salt concentration, single-channel insertions were observed (n ≥ 6) at an applied transmembrane potential of +80 mV and in 10 mM potassium phosphate, pH 7.4. **b,** Single-channel current traces acquired with FhuA ΔC/Δ4L protein nanopore at various KCl concentrations, and in 10 mM potassium phosphate, pH 7.4. **c**, FhuA ΔC/Δ4L pore inserts efficiently into a lipid bilayer at low (the top trace) and high (the bottom trace) KCl concentrations. **d**, Single-channel electrical traces acquired with a bilayer patch, which contained three FhuA ΔC/Δ4L protein nanopores. Data was collected at various temperatures and in 1 M KCl, 10 mM potassium phosphate, pH 7.4. The applied transmembrane potential was +80 mV. The electrical traces were low-pass Bessel filtered at 2 kHz.

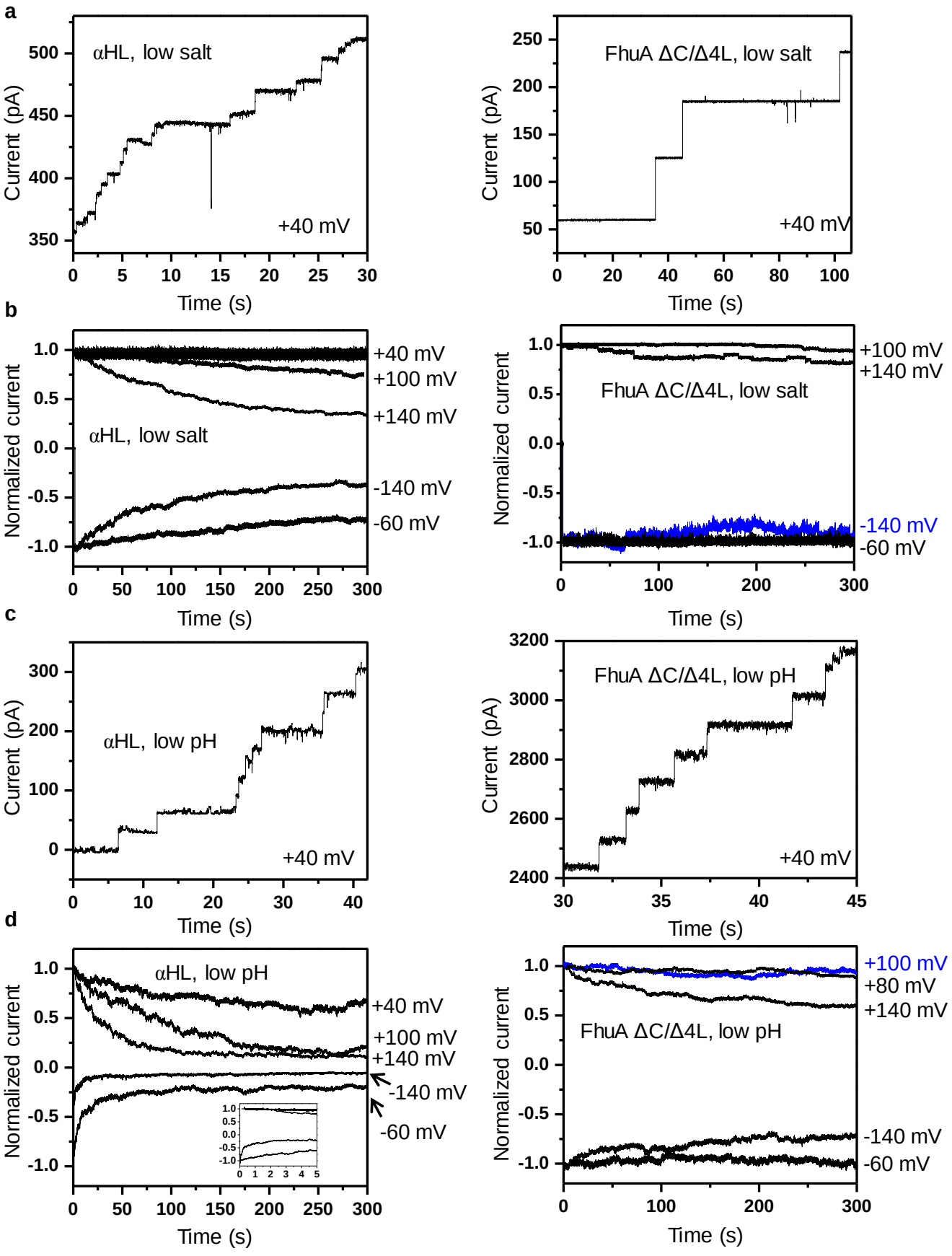

a
αHL, low salt
+40 mV
Current (pA)
Time (s)
FhuA ΔC/Δ4L, low salt
+40 mV
Current (pA)
Time (s)
b
αHL, low salt
+40 mV
+100 mV
+140 mV
-140 mV
-60 mV
Normalized current
Time (s)
FhuA ΔC/Δ4L, low salt
+100 mV
+140 mV
-140 mV
-60 mV
Normalized current
Time (s)
c
αHL, low pH
+40 mV
Current (pA)
Time (s)
FhuA ΔC/Δ4L, low pH
+40 mV
Current (pA)
Time (s)
d
αHL, low pH
+40 mV
+100 mV
+140 mV
-140 mV
-60 mV
Normalized current
Time (s)
FhuA ΔC/Δ4L, low pH
+100 mV
+80 mV
+140 mV
-140 mV
-60 mV
Normalized current
Time (s)

Fig. S6. Comparison of channel closure rates between α-HL and the engineered FhuA ΔC/Δ4L. **a**, Snapshots of stepwise channel insertions of αHL (left) and FhuA ΔC/Δ4L (right), just before the macroscopic recordings, in 200 mM KCl, 10 mM phosphate buffer, pH 7.4; **b**, Macroscopic current recordings collected with αHL (left) and the engineered FhuA ΔC/Δ4L (right) at various transmembrane potentials. The buffer conditions are the same as in **a**. **c**, Snapshots of stepwise channel insertions of αHL (left) and FhuA ΔC/Δ4L (right), just before the macroscopic recordings, in 1 M NaCl, 10 mM phosphate-citrate buffer, pH 3.5. **d**, Macroscopic current recordings collected with αHL (left) and the engineered FhuA ΔC/Δ4L (right) at various transmembrane potentials. The buffer conditions are the same as in **c**. Inset shows a high-resolution time scale with the very first closures of the αHL nanopores at a transmembrane potential of -140 mV. Numbers of protein nanopores in a bilayer patch varied from 50 to 100 at low salt concentration as in **b**, and 30 to 45 at acidic pH as in **d**. Electrical traces were low-pass Bessel filtered at 200 Hz in **a** and **c**.

Table S2. Closure rates (k) recorded with αHL and FhuA ΔC/Δ4L nanopores under low salt concentration (**a**) and acidic pH (**b**) conditions. For **a**, all experiments were carried out in 200 mM KCl, 10 mM potassium phosphate, pH 7.4. For **b**, all experiments were performed in 1 M NaCl, 10 mM phosphate-citrate buffer, pH 3.5. Numbers of nanopores were as in Fig. S6. Rates were derived from 15 minute-long traces. Closure rates were derived from fitting the macroscopic current trace with single exponential. Residual current percentage (I_r) was calculated from the ratio between the current at infinity (I_∞) and the initial current (I_0). The values represent means ± SDs calculated from three separate multi-channel survival curve experiments. NA= fitting was not possible with these traces as they represent almost straight horizontal lines.

a

	+100 mV		+140 mV		-60 mV		-140 mV	
	k (s^{-1}) x 10^{-3}	I_r (%)	k (s^{-1}) x 10^{-3}	I_r (%)	k (s^{-1}) x 10^{-3}	I_r (%)	k (s^{-1}) x 10^{-3}	I_r (%)
αHL	3 ± 0.9	75 ± 12	8 ± 3	50 ± 18	2.7 ± 1	60 ± 20	9 ± 2.5	45 ± 9
FhuA ΔC/Δ4L	NA	NA	3 ± 1	85 ± 10	NA	NA	NA	NA

b

	+100 mV		+140 mV		-60 mV		-140 mV	
	k (s^{-1}) x 10^{-3}	I_r (%)	k (s^{-1}) x 10^{-3}	I_r (%)	k (s^{-1}) x 10^{-3}	I_r (%)	k (s^{-1}) x 10^{-3}	I_r (%)
αHL	16 ± 3	20 ± 5	33 ± 10	15 ± 5	32 ± 10	20 ± 10	100 ± 30	10 ± 5
FhuA ΔC/Δ4L	NA	NA	3 ± 0.6	80 ± 10	NA	NA	4 ± 0.7	80 ± 15

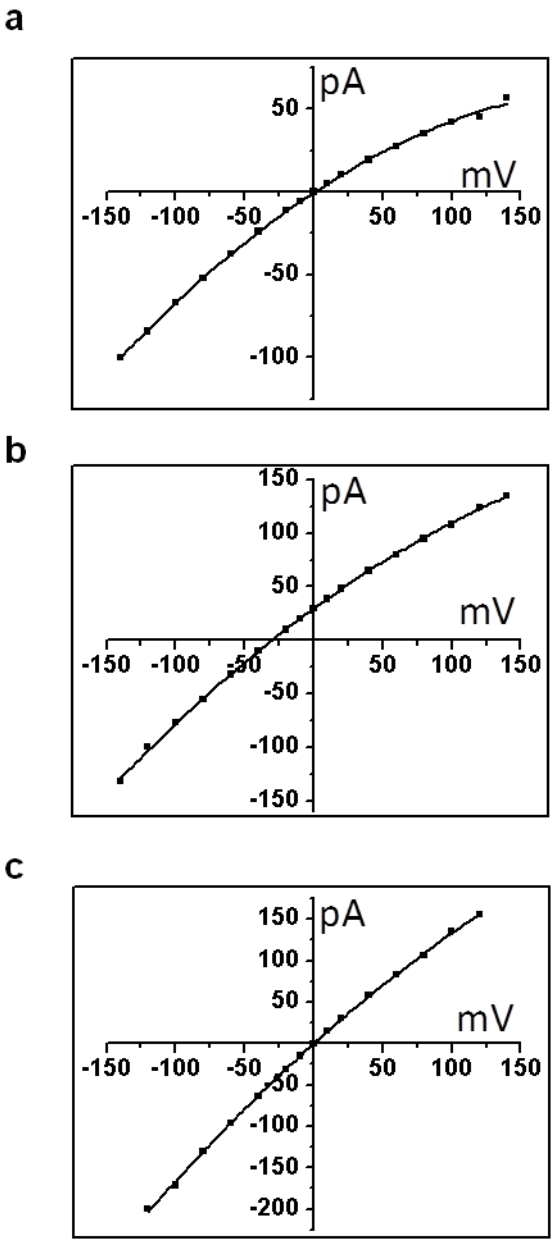


Fig. S7. The I-V curves of the single-channel currents through the FhuA ΔC/Δ4L protein nanopores in symmetrical solution of 0.02 M KCl, 10 mM potassium phosphate, pH 7.4 (**a**); asymmetric solutions of 0.02 M KCl, 10 mM potassium phosphate, pH 7.4 (*cis*) and 0.2 M KCl, 10 mM potassium phosphate, pH 7.4 (*trans*) (**b**); and in symmetrical solution of 0.2 M KCl, 10 mM potassium phosphate, pH 7.4 (**c**).

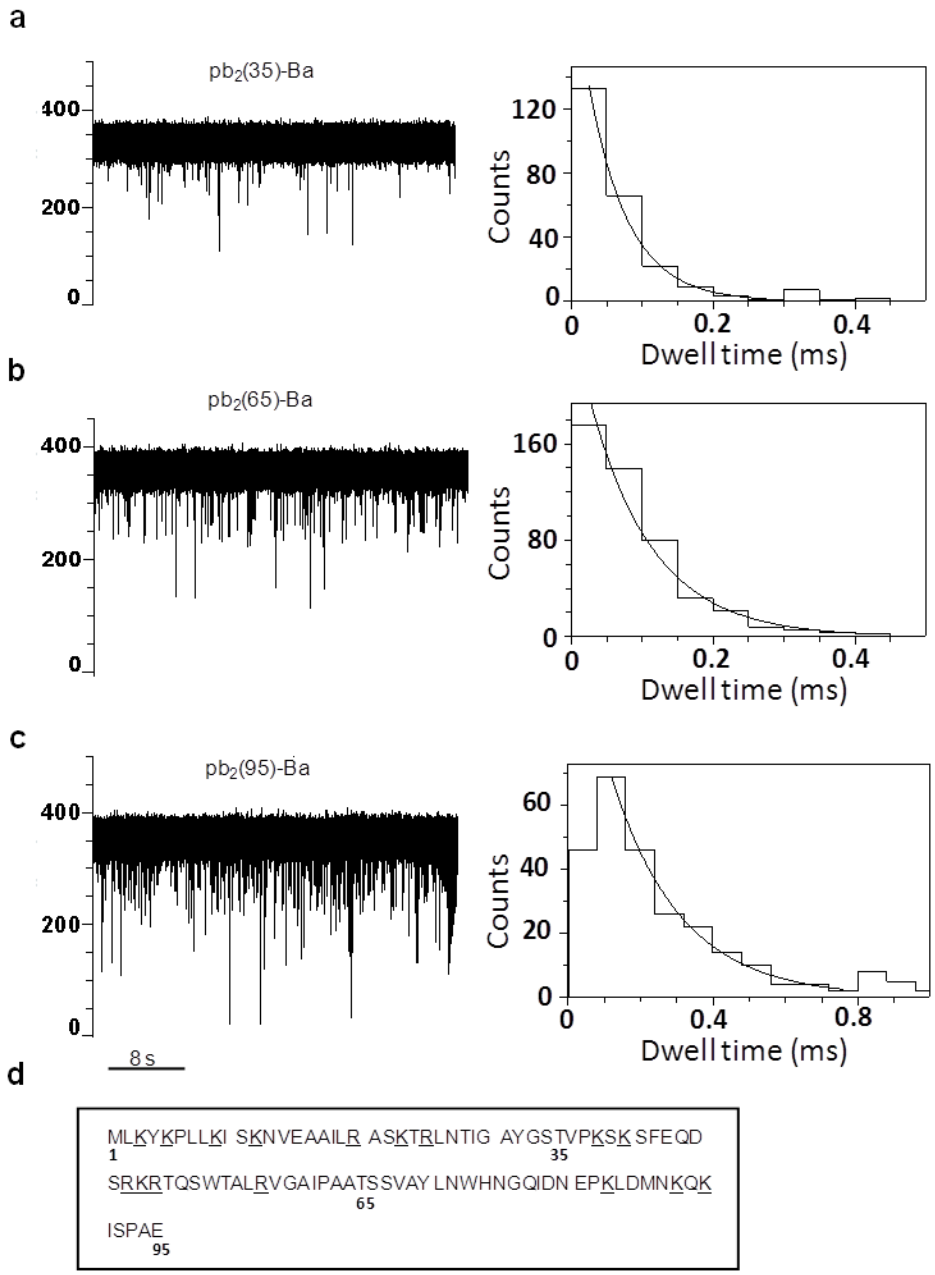


Fig. S8. Detection of modified proteins by the engineered FhuA ΔC/Δ4L protein nanopore. Single-channel electrical recordings performed with the engineered FhuA ΔC/Δ4L protein nanopore when pb_2(35)-Ba (**a**), pb_2(65)-Ba (**b**) and pb_2(95)-Ba (**c**) proteins were added to the *trans* chamber (0.2 μM). Right panels are representative histograms of the dwell times for each polypeptides. The smooth curves are single-exponential fits using log likelihood ratio tests. **d**, Primary amino acids sequences for the N-terminal region of pre-cytochrome b_2. Single-letter code is used to present the sequences and the underlined letters are the positively-charged amino acids at pH 7.4. 35, 65 and 95 numbers represent the length of pre-cytochrome b_2 (pb_2) that were used in constructing pre-cytochrome b_2-fused barnase polypeptides pb_2(35)-Ba, pb_2(65)-Ba and pb_2(95)-Ba, respectively. Single-channel recordings were performed at room temperature in 1 M KCl, 10 mM potassium phosphate, pH 7.4, and at a transmembrane potential of +80 mV. The single-channel electrical traces were low-pass Bessel filtered at 2 kHz.

Table S3: Characteristics of protein analytes and their interactions with the FhuA ΔC/Δ4L protein nanopore.

Analyte [a]	Number of residues [b]	Presequence charge [c]	Charge density [d]	k_{on} ($M^{-1}s^{-1}$) x 10^{-6} [e]	I_r (%) [f]
pb_2(35)-Ba	35	+7	0.2	11.5 ± 3.2	0.72 ± 0.18
pb_2(65)-Ba	65	+13	0.2	20.5 ± 3.3	0.55 ± 0.17
pb_2(95)-Ba	95	+16	0.17	37.5 ± 6.5	0.25 ± 0.12

[a]The complete sequences of the charged leading arms of the pb_2-Ba proteins are given in **Fig. S8**.
[b]The numbers of residues in the presequences pb_2(35)-Ba, pb_2(65)-Ba and pb_2(95)-Ba, respectively.
[c] The charge is estimated at pH 7.4.
[d]The charge density was calculated as the number of positively charged amino acids divided by the total number of amino acids.
[e]All experiments were carried out in symmetrical conditions of 1 M KCl, 10 mM potassium phosphate, pH 7.4. 200 nM of each protein was added to the *trans* chamber. The transmembrane potential was +80 mV. The values represent means ± SDs calculated from three distinct single-channel electrical recordings.
[f]The other conditions are the same as in the legend of **Fig. 4**. The values represent means ± SDs calculated from three distinct macroscopic current measurements.

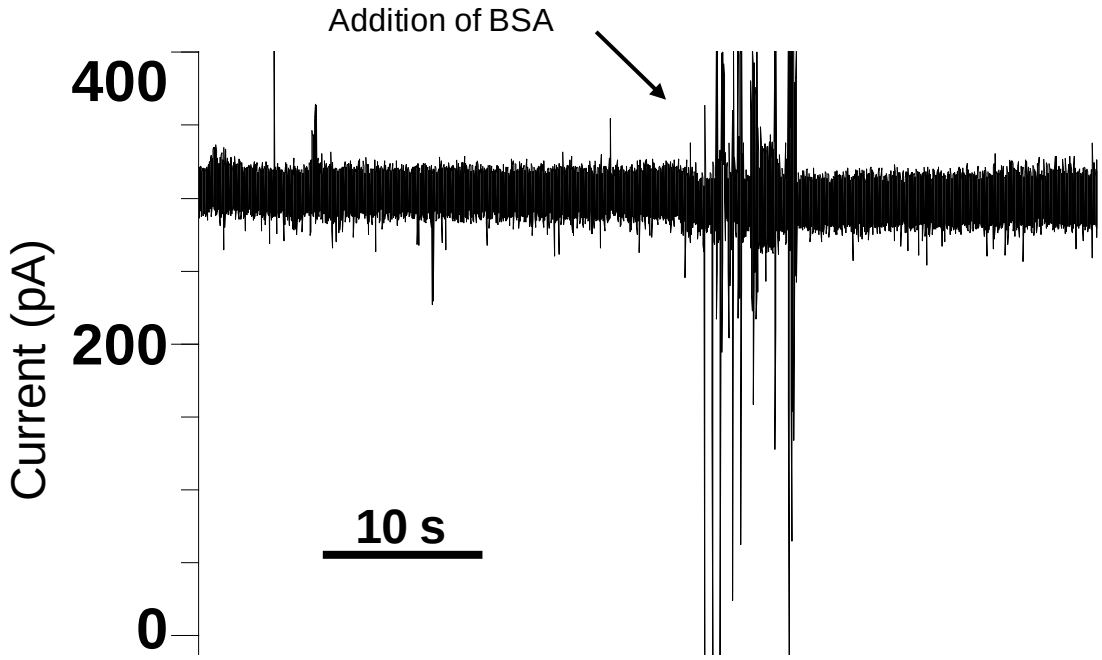


Fig. S9. Bovine serum albumin (BSA) does not interact with FhuA ΔC/Δ4L protein nanopore. Single-channel electrical recording of the FhuA ΔC/ Δ4L protein nanopore before and after the addition of BSA to the *trans* chamber (0.2 μM), and recorded in 1 M KCl, 10 mM potassium phosphate, pH 7.4. The transmembrane potential was +80 mV. The signal was low-pass Bessel filtered at 2 kHz.

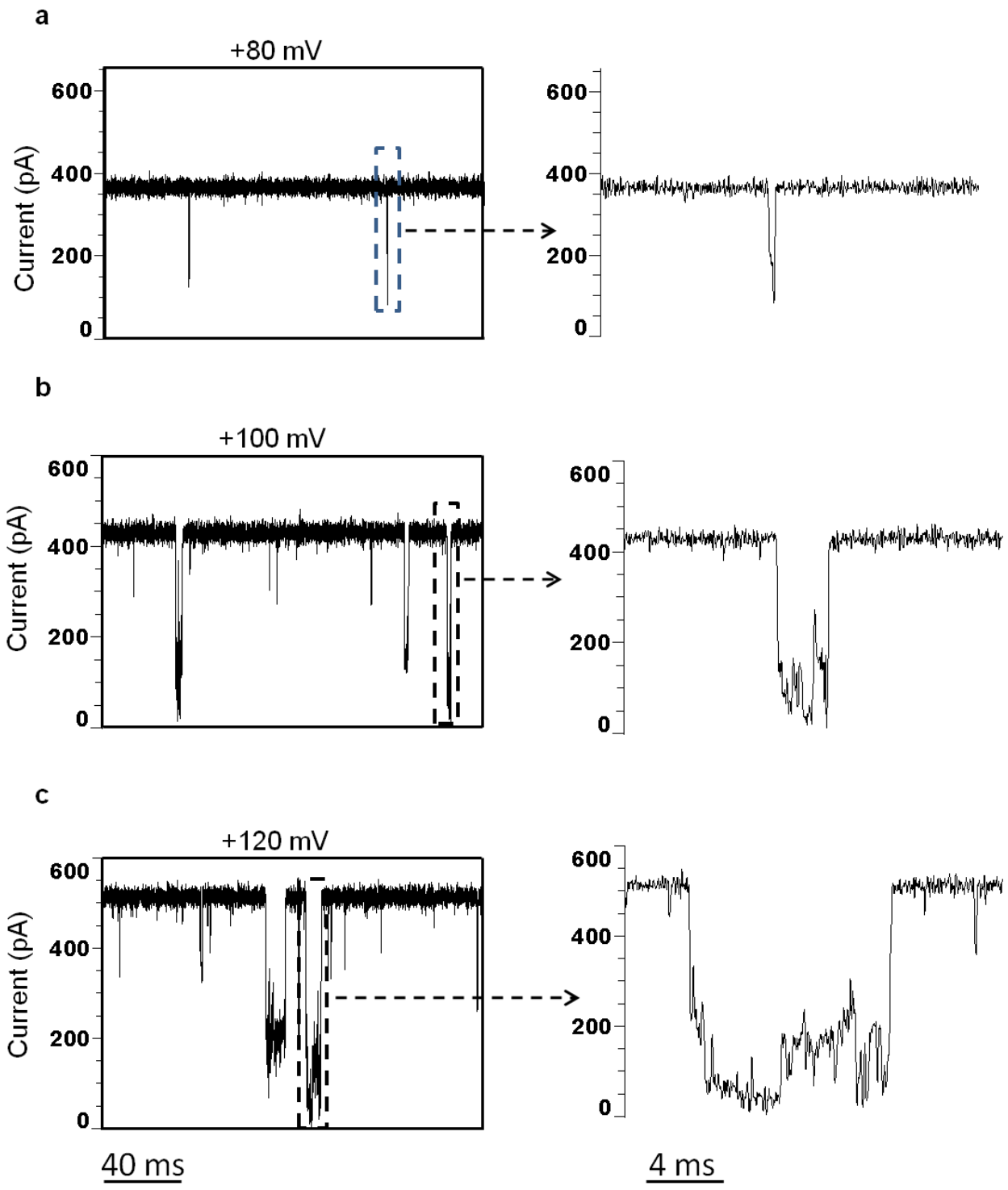


Fig. S10. Voltage dependence of the interaction between the pb_2(35)-Ba protein and the FhuA ΔC/Δ4L protein nanopore. Representative single-channel electrical recordings in the presence of the pb_2(35)-Ba (0.2 μM) at transmembrane potential of (**a**) +80 mV, (**b**) +100 mV and **(c**) +120 mV. Right panels are expanded traces from single events. The other experimental conditions of the single-channel recordings were the same as those in **Fig. S8**.

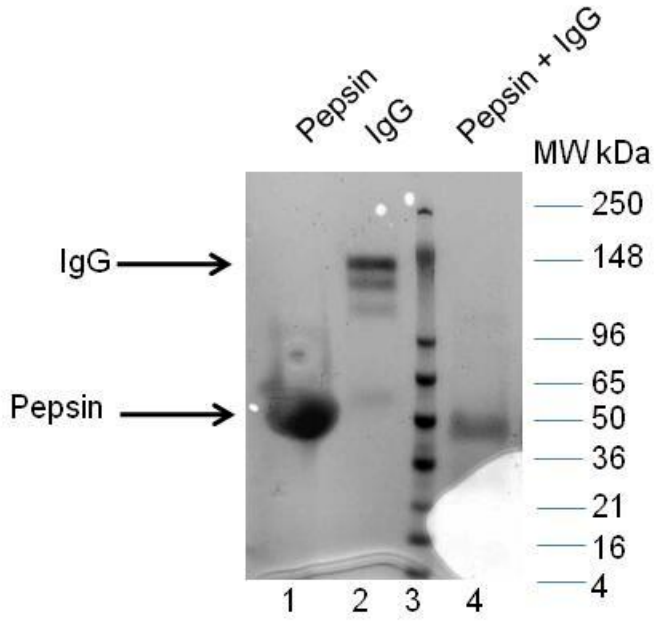


Fig. S11. Digestion of immunoglobulin G (IgG) by pepsin. Lane 1, 180 μg of pepsin (684 units); lane 2, 15 μg IgG (~ 1 pmole); lane 3, molecular weight standard (MW); Lane 4, 1 pmole of IgG was incubated with 1.14 units/μl of pepsin in 25 μl of 1 M NaCl, phosphate-citrate buffer, pH 3.9. All samples were quenched after 1.5 hours by the addition of 25 μl of 2 M Tris, pH 11.5. Samples were loaded and visualized as in **Fig. 1c**. MW stands for the molecular weight standard**.**

Table S4. Characteristics of the peptides resulted from the pepsin digestion of a single mouse IgG heavy chain[a]

Digestion site	Peptide[b]	Length (residues)	Mass (Da)	Isoelectric point[c]	Charge[d]	Hydropathy Index[e]
6	MGWTWI	6	793.0	6	0	3.5
22	SVTTGVHSDVQ	11	1129.2	4.9	1	-1.0
30	QQSGPEL	7	757.8	4	0	-9.5
45	EKPGASVKISCKASG	15	1461.7	9.8	+3	
63	PGHNINWIVQRNGKS	15	1719.9	11.5	+3	-16.5
78	LEWIGNIDPYYGGTN	15	1711.8	3	-1	-8.2
83	FNPKF	5	651.8	10.1	+1	-3.4
89	KGKATL	6	616.8	10.6	+3	-3.3
97	TVDKSSST	8	823.9	6.7	+1	-7.0
124	QSEDSAVYYCARRRDGNYG	19	2210.3	6.4	+2	-27.0
132	TYWGQGT	7	811.8	5.9	0	-7.9
148	VTVSAAKTTPPSVYP	15	1517.7	9.7	+1	2.5
162	LAPGSAAQTNSMVT	14	1347.5	6	0	3.3
183	PEPVTVTWNSGS	12	1273.4	3.3	0	-6.1
194	SSGVHTFPAV	10	1001.1	7.8	+1	5.5
255	SSSVTVPSSTWPSETVTCNVA HPASSTKVDKKIVPRDCGCK PCICTVPEVSSV	53	5506.3	7.7	+4	-1.0
266	IFPPKPKDVL	10	1153.4	9.9	+1	-0.8
289	LTPKVTCVVVDISKDDPEVQ	20	2185.5	4	0	1.1
310	VDDVEVHTAQTQPREEQ	17	1981.1	3.9	0	-24.3
328	PIMHQDW	7	926.1	4.9	+1	-6.3
333	NGKE	4	446.5			
343	FKCRVNSAAF	10	1142.3	10.1	+2	3.2
380	PAPIEKTISKTKGRPKAPQVY TIPPPKEQMAKDKVSL	37	4073.9	10.5	+7	-32.9
419	PEDITVEWQWNGQPAENYK NTQPIMNTNGSY	31	3625.9	3.7	0	-41.9
437	NVQKSNWEAGNT	12	1347.4	6.0	0	-18.2
455	HNHHTEKS	8	989.0	8.1	+3	-22.0
482	LSHSPGKPQVYTIPPPKEQMA KDKVSL	27	2976.5	10	+4	-21.4
521	PEDITVEWQWNGQPAENYK NTQPIMNTNGSY	31	3625.8	4	0	-42.2

[a]Predicted digestion products were obtained using ExPASy Peptide Cutter Tool [18]. This table presents the digestion products of a single mouse IgG heavy chain [19]**.** The digestion was set at pH>2, assuming no protein modifications, and that any potential cysteine disulfide bonds were reduced. [b]Peptides that are shorter than 5 residues in length are not displayed. [c]The isoelectric point for each peptide was obtained from ExPASy PortParam Tool [18]. [d]The charge is estimated at pH 3.9, which is identical to the

digestion condition in the nanopore recordings. [e]Kyte-Doolittel method was used to calculate the values of the hydrophathy index [20]. These values are the sums of the hydropathy indexes for the individual amino acids in the predicted peptide. The positive and negative values indicate the hydrophobic and the hydrophilic nature of the peptide, respectively.

Table S5. Characteristics resulted from the pepsin digestion of a single mouse IgG light chain[a]

Digestion site	Peptide	Length (residues)	Mass (Da)	Isoelectric point	Charge	Hydropathy index
3	TQSPAT	6	603.6	5.2	0	-5.5
20	SVTPGDSVS	9	847.8	3.8	0	-0.2
32	SCRASQSISNN	11	1166.2	7.9	+1	-9.4
46	HWYQQKSHESPRL	13	1695.8	8.6	+2	-26.9
62	LIKYTSQSMSGIPSRF	16	1815.1	10	+2	-1.6
70	SGSGSGTD	8	666.6	3.8	0	-7.8
82	SINSVETED	9	992.9	3.6	0	-7.6
97	CQQSGSWPRT	10	1149.3	8.3	+1	-14.2
103	FGGGTK	6	565.6	8.7	+1	-3.0
124	DIKRADAAPTVSIFPPSSEQ	20	2129.4	4.6	+2	-8.9
134	TSGGASVVC	9	779.8	5.8	0	9.6
159	FYPKDINVKWKIDGSERQNGV	21	2493.8	8.4	+3	-22.4
178	NSWTDQDSKDSTYSMSST	18	2040.1	3.9	+1	-28.6

[a]This table presents the digestion products of a single mouse IgG light chain [21].

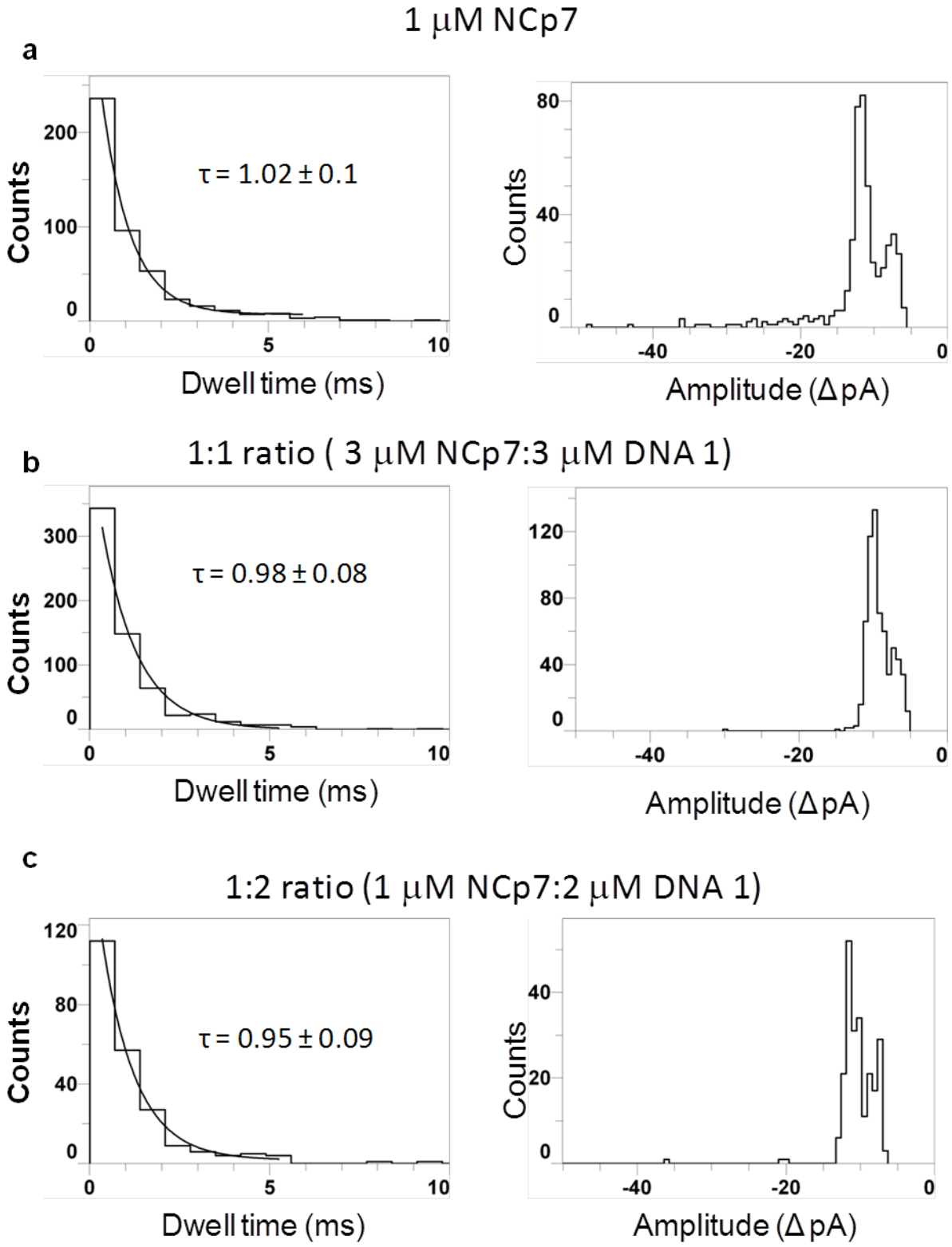


Fig. S12. Characteristics of the NCp7-produced, single-channel current blockades observed with FhuA ΔC/Δ4L nanopore. a, Data was recorded prior to the addition of the DNA 1 aptamer. **b**, Data was recorded when 1:1 ratio of NCp7: DNA 1 was present in the *trans* side of chamber; **(C)** Data was recorded when 1:2 ratio of NCp7: DNA 1 was present in the *trans* side of chamber. The left panels are

representative histograms of the event dwell times. The smooth curves are single-exponential fits using the log likelihood ratio (LLR) tests [22,23]. The right panels are representative current amplitude histograms. Single-channel electrical recordings and buffer conditions were the same as in **Fig. 5d**.

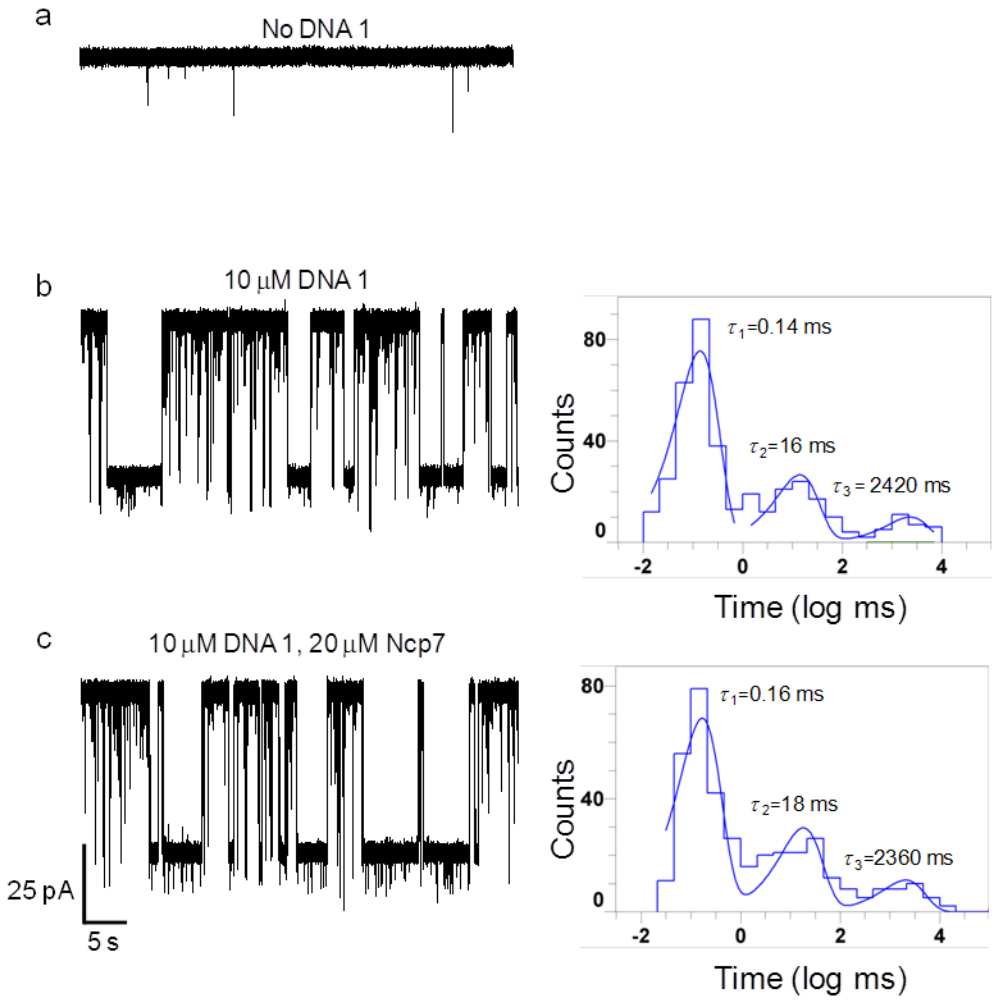


Fig. S13. Exploration of the kinetics of Ncp7-DNA 1 interactions using the αHL nanopore at 500 mM KCl. Representative single channel electrical recordings for (**a**) αHL nanopore, (**b**) αHL nanopore with DNA 1 (10) μM , (**c**) αHL nanopore with DNA 1 (10 μM) and Ncp7 (20 μM). αHL nanopores were inserted from the *cis* side of the chamber, whereas DNA 1 and Ncp7 protein were also added to the *cis* side. Single-channel electrical recordings were performed at a transmembrane potential of +120 mV and with the buffer solution containing 500 mM NaCl, 5 mM potassium phosphate, pH. 7.4. Right panels in (**b**) and (**c**) are semi-logarithmic histograms for dwell time analysis. The dwell times were best fitted with three-exponential time distribution. In **b**, the event frequencies for τ_1, τ_2 and τ_3 were 0.73, 0.31 and 0.12 s^{-1}, respectively. In **c**, the event frequencies for τ_1, τ_2 and τ_3 were 0.74, 0.32 and 0.10 s^{-1}, respectively. No major alterations of frequency and dwell time of events were observed in **b** and **c**. All single-channel electrical traces were low-pass Bessel filtered at 2 kHz.

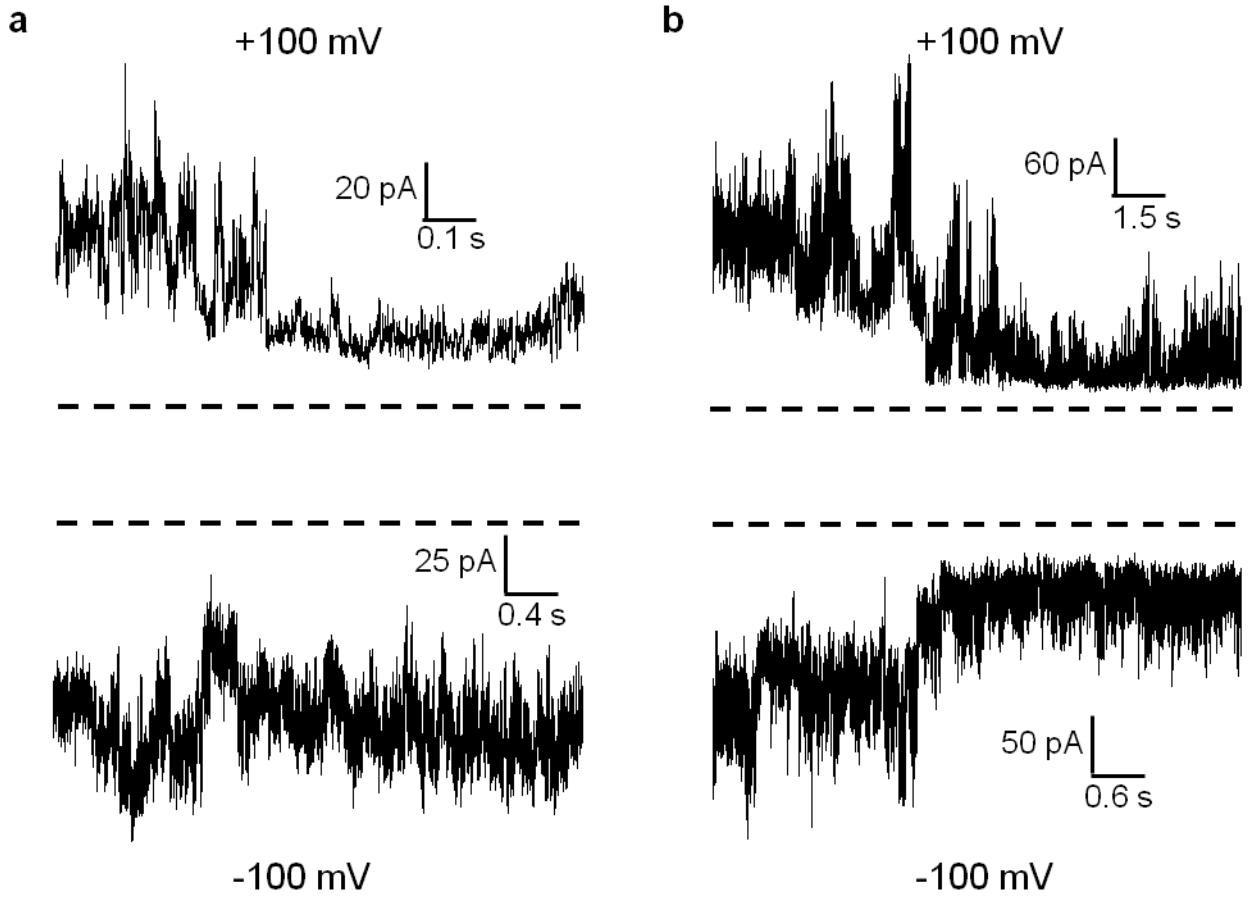


Fig. S14. Sensitivity of membrane-extracted FhuA ΔC/Δ4L to low salt concentration or highly acidic pH. **a,** Representative single-channel recordings with membrane-extracted FhuA ΔC/Δ4L in 200 mM KCl, 10 mM potassium phosphate, pH 7.4. The protein nanopore either closed to ~ 30% of its maximum open current state with preceding current fluctuations that range from 10% to 30% when a transmembrane potential of +100 mV was applied (the top panel), or showed fluctuations in the current baseline when a transmembrane potential of -100 mV was applied (the bottom panel). **b,** Representative single-channel recordings with membrane-extracted FhuA ΔC/Δ4L in 1 M NaCl, 10 mM phosphate-citrate, pH 3.5. The protein nanopore closed to ~ 20% of its maximum open current state with preceding fluctuations in the current baseline when a transmembrane potential of +100 mV (the top panel), or -100 mV (the bottom panel) was applied. FhuA ΔC/Δ4L protein was prepared as previously described [1]. Dashed lines represented the zero current. All electrical traces were low-pass Bessel filtered at 2 kHz.

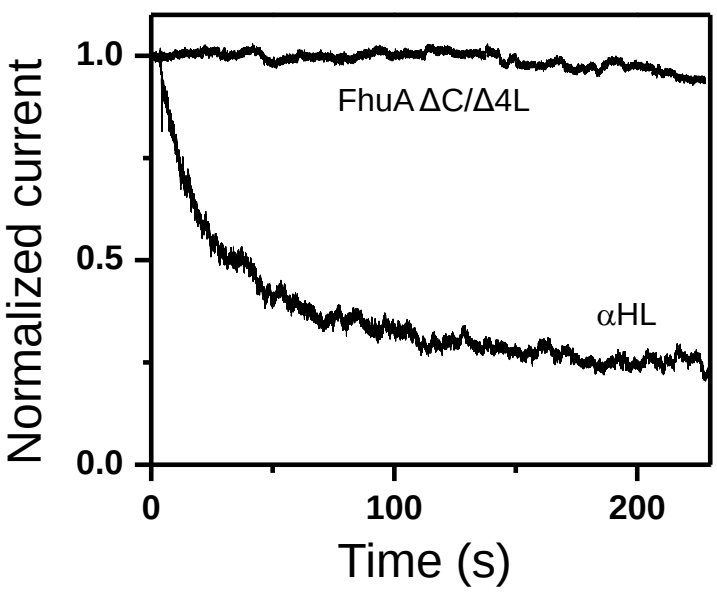


Fig. S15. Comparison between αHL and the engineered FhuA ΔC/Δ4L nanopores at low physiological salt concentration and highly acidic pH. This panel illustrates the multi-channel survival curves recoded with each nanopore in 150 mM NaCl, 10 mM phosphate-citrate buffer, pH 3.5. The applied transmembrane potential was +100 mV. No detectable insertions of channels were observed for several minutes before the onset of the macroscopic current recordings. The curves presented in this panel represent the changes of a total multi-channel current given by either ~120 αHL nanopores or ~100 FhuA ΔC/Δ4L nanopores.

Reference List

1. Mohammad, M. M., Howard, K. R. & Movileanu, L. Redesign of a plugged beta-barrel membrane protein. *J. Biol. Chem.* **286**, 8000-8013 (2011).

2. Endriss, F. & Braun, V. Loop deletions indicate regions important for FhuA transport and receptor functions in Escherichia coli. *J. Bacteriol.* **186**, 4818-4823 (2004).

3. Kleinschmidt, J. H. & Tamm, L. K. Time-resolved distance determination by tryptophan fluorescence quenching: probing intermediates in membrane protein folding. *Biochemistry* **38**, 4996-5005 (1999).

4. Mohammad, M. M., Prakash, S., Matouschek, A. & Movileanu, L. Controlling a single protein in a nanopore through electrostatic traps. *J. Am. Chem. Soc.* **130**, 4081-4088 (2008).

5. Mohammad, M. M. & Movileanu, L. Excursion of a single polypeptide into a protein pore: simple physics, but complicated biology. *Eur. Biophys. J.* **37**, 913-925 (2008).

6. Paoletti, A. C., Shubsda, M. F., Hudson, B. S. & Borer, P. N. Affinities of the nucleocapsid protein for variants of SL3 RNA in HIV-1. *Biochemistry* **41**, 15423-15428 (2002).

7. Shubsda, M. F., Paoletti, A. C., Hudson, B. S. & Borer, P. N. Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein. *Biochemistry* **41**, 5276-5282 (2002).

8. Athavale, S. S., Ouyang, W., McPike, M. P., Hudson, B. S. & Borer, P. N. Effects of the nature and concentration of salt on the interaction of the HIV-1 nucleocapsid protein with SL3 RNA. *Biochemistry* **49**, 3525-3533 (2010).

9. Athavale, S. S. Affinity of the HIV-1 Nucleocapsid Protein for SL3 RNA: Effects Due to Salt and Variation Due to Measurement Technique. 2010.
Ref Type: Thesis/Dissertation

10. Song, L. Z. *et al.* Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science* **274**, 1859-1866 (1996).

11. Simpson, A. A. *et al.* Structure of the bacteriophage phi29 DNA packaging motor. *Nature* **408**, 745-750 (2000).

12. Faller, M., Niederweis, M. & Schulz, G. E. The structure of a mycobacterial outer-membrane channel. *Science* **303**, 1189-1192 (2004).

13. Subbarao, G. V. & van den Berg, B. Crystal structure of the monomeric porin OmpG. *J. Mol. Biol* **360**, 750-759 (2006).

14. Locher, K. P. *et al.* Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell* **95**, 771-778 (1998).

15. Bezrukov, S. M. & Vodyanoy, I. Probing alamethicin channels with water-soluble polymers - Effect on conductance of channel states. *Biophys. J.* **64**, 16-25 (1993).

16. Bezrukov, S. M., Vodyanoy, I., Brutyan, R. A. & Kasianowicz, J. J. Dynamics and free energy of polymers partitioning into a nanoscale pore. *Macromolecules* **29**, 8517-8522 (1996).

17. Hille, B. *Ion Channels of Excitable Membranes*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA (2001).

18. Gasteiger, E. *et al. Proteomics Protocols Handbook*. Walker J.M. (ed.), pp. 571-607 (Humana Press,2005).

19. Bowdish, K., Tang, Y., Hicks, J. B. & Hilvert, D. Yeast expression of a catalytic antibody with chorismate mutase activity. *J. Biol. Chem.* **266**, 11901-11908 (1991).

20. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132 (1982).

21. Cauerhff, A., Goldbaum, F. A. & Braden, B. C. Structural mechanism for affinity maturation of an anti-lysozyme antibody. *Proc. Natl. Acad. Sci. U. S. A* **101**, 3539-3544 (2004).

22. McManus, O. B., Blatz, A. L. & Magleby, K. L. Sampling, Log Binning, Fitting, and Plotting Durations of Open and Shut Intervals From Single Channels and the Effects of Noise. *Pflugers Arch.* **410**, 530-553 (1987).

23. McManus, O. B. & Magleby, K. L. Kinetic states and modes of single large-conductance calcium-activated potassium channels in cultured rat skeletal-muscle. *J. Physiol. (Lond.)* **402**, 79-120 (1988).